DOCUMENT RESUME

ED 302 064 FL 017 687

AUTHOR Echols, Catharine H.
TITLE The Role of Stress, Position and Intonation in the Representation and Identification of Early Words.
PUB DATE Aug 88
NOTE 9p.; In: Papers and Reports on Child Language Development, Volume 27; see FL 017 572.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Child Language; *Intonation; *Language Acquisition; Language Processing; Language Research; Listening Comprehension; *Stress (Phonology); *Syllables; Toddlers; *Verbal Learning; Word Recognition

ABSTRACT

Two studies examined children's perceptual biases in extracting or identifying words from the stream of speech. In one study, evidence for the salience of stressed and final syllables was found. Young children less frequently omitted those syllables from their productions and produced unstressed and nonfinal syllables less accurately. A second study explored the possibility that attention to the stress pattern, or overall intonation contour, of a speech sequence may assist in children's identification of words. Where subjects were required to identify referents for words on the basis of either stress pattern or phoneme sequence, 3-year-olds were more likely than adults to make choices based on intonational similarity. These observations suggest that children are more likely than adults to include intonation as an important part of the representation for a word. They are also at least consistent with the view that stress pattern may be extracted as part of the representation of a word and may, in some cases, serve to define a word for a child. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED302064

PRCLD 27 (1988)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E Clark
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

THE ROLE OF STRESS, POSITION AND INTONATION IN THE REPRESENTATION AND IDENTIFICATION OF EARLY WORDS

Catharine H. Echols
Stanford University

The question of how children initially extract and represent approximately word-sized units from the stream of adult speech is one of the most fundamental of language acquisition. Unless the child can identify word-level units, virtually all linguistic accomplishments will be impossible. It is also a far more difficult problem than it may initially appear: Contrary to adult intuitions, words are rarely separated by pauses and there are generally few, if any, consistent cues to boundaries between words (Cole & Jakamik, 1980; Hayes & Clark, 1970).

It has been suggested that certain attentional biases of the child may limit the difficulty of this task. Rather than attending to and attempting to segment an entire sequence of adult speech, the very young language learner may be predisposed to focus primarily on certain elements of that speech sequence. Those elements would tend to be extracted and should be prominent in the child's early representations and productions.

The notion that certain elements of the input speech are particularly salient to young language learners was originally proposed by Slobin (1973) in the context of the acquisition of morphology. On the basis of cross-linguistic data on the order of acquisition of various morphological inflections, Slobin argued, for example, that children pay particular attention to elements at the ends of utterances. More recently, perceptual biases have also been discussed in the context of segmentation. Thus, for example, Gleitman and Wanner (1982) have argued that a tendency to attend to and extract stressed syllables will account for the initial segmentation of speech.

If children are, in fact, biased to attend to stressed and final syllables and if those biases do assist the child in identifying approximately word-level units in the stream of speech, then those perceptual biases should be reflected in first utterances. More specifically, first utterances should tend to contain primarily stressed and final syllables. In fact, stress and position do appear to account for many first utterances. Thus, for example, children may produce /čak/ for "chocolate," /lai/ for "butterfly," or /rai-sə/ for "eraser," each of which contain only stressed, final or stressed and final syllables.

Further support for this position comes from experimental and descriptive studies of children's imitations, as well as some crosslinguistic data. Children will tend to include stressed or final syllables in their productions even for imitations of nonsense-words (Blasdell & Jensen, 1970; Frumhoff & Newport, in preparation) and for languages in which stress falls on inflectional morphemes rather than on the content words of the language (Fueur, 1980; Pye, 1983).

An important test of the claim that the proposed biases can account for the nature of first words could be provided by a systematic analysis of a corpus of early utterances. I will, therefore, give a brief description of a study which was intended to do just that. As will be seen, however, biases to attend to stressed and final syllables, while accounting for a great many first utterances, can by no means account for all. It has, in fact, been argued, that children may extract, as single units, longer sequences of speech which are defined by intonation or rhythm (Peters, 1977, 1983). Furthermore, although children's productions may be suggestive as to which elements of speech are perceived as most salient, productions are, of course, not

FL017687

necessarily identical to representations. With respect to the former of these concerns, the second study, and that which is the primary focus of this paper, investigated the possibility that prosodic properties may be important in the early extraction and representation of word-like elements of speech. In the interest of assessing children's representations more directly, the data for this study were children's recognitions rather than their productions.

Analysis of Early Utterances

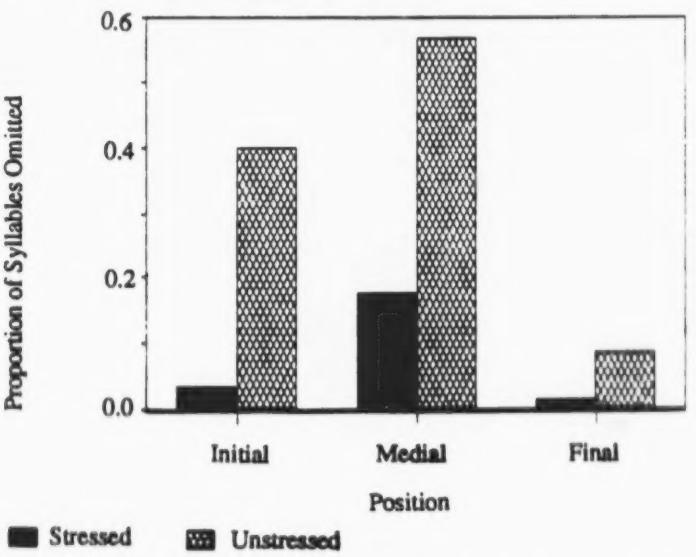
The corpus submitted to analysis was collected from three children between the ages of 17 and 23 months, each of whom was producing primarily single-word utterances. Children were tape-recorded in their home during natural interactions between the child and parent or child and experimenter. Utterances with adult targets longer than a single syllable (a total of 615 utterances) were phonetically transcribed and coded in various ways. For present purposes, certain syllable-level codings are most relevant: Each syllable of the child's utterance was coded, in relation to the corresponding syllable of the adult target, for presence or absence and for accuracy.

Two primary predictions were tested. The first was that syllables which were unstressed and nonfinal should more frequently be omitted than stressed or final syllables. Secondly, because stressed and final syllables are particularly salient, they should be more completely extracted. As a result, where unstressed syllables are included in a child's production, they should be less accurately produced than stressed or final syllables.

Both predictions were borne out. Figure 1 shows the proportion of syllables omitted as a function of stress and position. As can be seen, syllables which were both unstressed and

Figure 1

Proportion of Syllables Omitted as a Function of Position and Stress

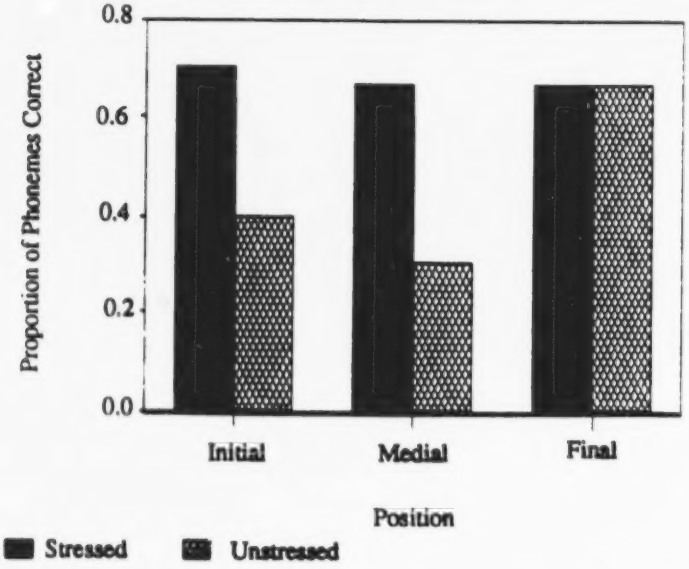


nonfinal were omitted far more frequently than stressed or final syllables. In Figure 2, the

proportion of phonemes correct is shown as a function of the stress level and position of the target syllable. Syllables for which the target was stressed or final were also produced more accurately

Figure 2

Proportion of Phonemes Correct as a Function of Position and Stress



than unstressed or nonfinal syllables. These results are consistent with the position that young language learners are particularly likely to focus on and extract stressed and final syllables.

Children do not, however, extract only stressed and final syllables. Children will frequently produce utterances containing filler or "dummy" syllables, in which syllables of the adult target have been reduced to a schwa or other undifferentiated sound (e.g., /m-bɛ-dɛ/ for "all-better" or /ʌ-tat/ for "what's that"). These filler syllables may frequently correspond to target syllables which are neither stressed nor final. In such cases, the child has clearly stored some representation for the unstressed, nonfinal syllable but it is a representation which appears to lack the phonemes of the target syllable. Children also produce utterances which include only some of the phonemes of the adult target, and which appear to be defined primarily by the stress pattern or intonation of the adult target (e.g., /da-dʌ-do/ for "here-we-go," /hi-i-ya/ for "here-they-are"). I will describe such utterances as "underanalyzed sequences," although they have also been described by Peters (who identified the phenomenon in a 1977 paper) as "tunes" and "gestalt" speech. Filler syllables and underanalyzed sequences make evident the need for something more than only biases to attend to stressed and final syllables in a model of first utterances.

Turning back to stress for a moment, the research described earlier indicates that stress has an effect on what is extracted. The existence of these other types of utterances suggest another potential role for stress: A child may extract the stress pattern, itself, or stress as it contributes to the overall intonation contour or rhythm of an utterance. If a perspective is taken within which the child may extract a sequence defined by rhythm, intonation or stress independently of the

phonemes, then filler syllables could, for example, be described as the extraction of a rhythmic syllable without the associated phonemes. Underanalyzed sequences would then be described as the extraction of the stress or pitch pattern of a word or phrase, with only a few of the associated phonemes.

An extreme test of the view that children may attend to the stress *pattern*, and not only to stressed syllables, would be to construct a situation in which children must choose between stress pattern and phoneme sequence as the basis for identifying the referent for a newly learned label. In that way, it would be possible to assess whether, in some cases, the stress pattern may be more prominent in the child's representation of a word than the segmentals. This second study was designed for that purpose. It should be mentioned here, however, that although I would like to argue that it can be intonation, or one of several components of the prosodic contour of a speech sequence, that children will attend to and extract, the study described here tests for a tendency to attend to only one such component: the stress pattern. Additional research will be required to determine which components of an overall intonation contour a child may attend to, extract, and store as part of the representation for a word.

Method

Subjects. Subjects were 32 two-year olds and 32 three-year olds, from five daycare centers located in several Midwestern communities. Half of the children in each age group were assigned to an experimental condition and half to a control. In addition, 16 adults, undergraduates at the University of Illinois, participated in the experimental condition.

Design and Procedure. Children were taught names for two objects, one at a time. The names were presented in carrier sentences, such as, "that's a /buta*ne*/" (where italics indicate primary stress). The child was asked to repeat the object names and was then presented with both objects and tested to be sure he or she knew the names.[1] Once the child had twice successfully identified each object (in response to "where's the .."), the child was presented with the test stimulus--a word which carried the stress pattern of one of the trained labels but the segmental sequence of the other. The child was asked, as before, to choose the object corresponding to that name. For example, a child who had been taught /buta*ne*/ for one object and /wo*l*æ zi/ for the second object, might be asked "where's the /wo*l*æ zi/?" The dependent measure was whether the subject chose the object that had previously been labelled with the intonationally or the segmentally similar name. Responses in which the child chose the object which had been associated with the intonationally similar label were coded as "intonational" while those in which the choice was the object associated with the segmentally similar label were coded as "segmental." A total of three sets of two object labels, varying in the degree of similarity between the two words of each pair, were taught to each subject. The word pairs and stress patterns are shown in Table 1. The stress pattern assigned to each word, and the particular word serving as the test, were fully counterbalanced between subjects.

[1] Although it would have been preferable, for consistency purposes, to have the labelling taught by a tape, piloting indicated that children could not learn the words from the tape. Object names were, therefore, taught by an experimenter. Acoustic comparisons of a subset of the stimulus utterances are being carried out to assure that there are no differences in the stimuli presented to subjects of different ages or in different conditions. Although those analyses are not completed, preliminary comparisons suggest that there are no important differences.

Table 1

Wordsets and Stress Patterns for Study 2

Wordset	Wordpair	Stress Pattern
Wordset 1	kibamu gidanu	0X0 X00
Wordset 2	togufa dopæ sa	X00 00X
Wordset 3	wolæ zi butane	00X 0X0

Note. X = primary stress.

The procedure for children in the control condition was the same as that for those in the experimental condition except that the pair of words heard by each child did not differ in intonation, and the test word was identical to one of the trained words. The word pairs and contours were those used in the experimental condition and were, again, fully counterbalanced between subjects.

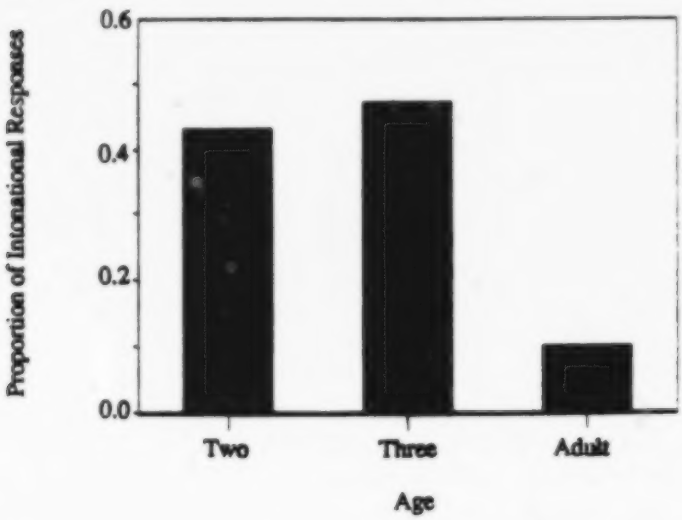
The proportion of "incorrect" responses obtained in the control condition served as a baseline to which the proportion of "intonational" responses obtained in the experimental condition could be compared. The justification for this is that incorrect responses reflect instances in which, in the absence of conflicting information, the child makes a "non-segmental" response. It was necessary to use such a baseline for the comparison, rather than comparing the proportion of intonational responses to chance (50%), because the claim is not that children will *always* attend to intonation rather than phoneme sequence. My previous work clearly shows that they frequently do attend to and extract a stressed or final syllable and that those syllables are frequently analyzed quite accurately. What I want to argue is that the stress pattern or other intonational characteristics may, in some cases, be extracted in lieu of, or along with, the phonemes of the stressed and final syllables and that these intonational characteristics will more frequently be prominent in the child's representation of a word than in the adult's. As a result, the stress pattern or intonation may sometimes override the segmental characteristics of a word in determining the child's perception of similarity between words. For adults, in contrast, the phoneme sequence should be far more prominent in the representation of a word than the stress pattern. Thus, adults should virtually always perceive similarity on the basis of segmental similarity rather than intonational similarity.

Results. A score reflecting the number of intonational responses was calculated for each child. Figure 3 shows the proportion of intonational responses for experimental subjects in each of the three age groups. Because each of the three wordsets were different in nature, the validity of a partial score was questionable. Any subjects who had produced one or more uncodable responses were therefore omitted from the analysis. (Two three-year olds and four two-year olds were omitted from analyses for this reason.) Chi-square analyses revealed a difference between children and adults in proportion of intonational responses, with children choosing the object associated with the intonationally similar label far more frequently than adults, $\chi^2(3) = 17.4$, $p < .001$. Two-year olds did not, however, differ from three-year olds, $\chi^2(3) = 1.37$, $p > .25$.

To determine the meaningfulness of these results, it is necessary to compare the pattern of intonational responses to the baseline obtained from children participating in the control

Figure 3

Proportion of Intonational Responses Across Three Age Levels



condition. Table 2 permits comparisons of the proportion of intonational responses for children and adults in the experimental condition and the proportion of incorrect responses for children in the control condition. Two-year olds, in particular, show fairly high error rates. As a result,

Table 2

Proportion of Intonational/Incorrect Responses as a Function of Age and Condition

	Age		
Condition	Two	Three	Adult
Experimental	.43	.47	.10
Control	.31	.23	---

Note. Scores for Experimental condition reflect intonational responses; those for Control condition reflect incorrect responses.

although the proportion of intonational responses produced by three-year olds was significantly different from the proportion of errors produced in the control condition, $\chi^2(3) = 9.31, p < .05$, this was not true for two-year olds, $\chi^2(3) = 4.23, p > .2$.

Discussion. As predicted, children showed more intonationally-defined responses than did adults. Because, however, the intonational responses of two-year olds did not differ significantly from the baseline level of incorrect responses produced by children in the control condition, no

claims can be made concerning what the two-year olds were attending to. Three-year olds do, however, appear to judge similarity on the basis of stress pattern rather than phoneme sequence in some subset of the instances. That observation suggests that prosodic components of a speech sequence can, in some cases, be more prominent in the child's representation of a word than the phonemes of that sequence, and it is consistent with the claim that children may extract and store the stress pattern of a target utterance without a complete representation of the phonemes.

Although three-year olds and many two-year olds are past the initial phases of language learning, in which a tendency to attend to and extract intonationally-defined sequences should be most important, these observations are not irrelevant for understanding the earlier processes: A tendency important in the earliest stages of language acquisition would presumably not disappear immediately. Three-year olds should still show a greater tendency than adults to extract and store intonation as part of the representation of the form of a word. Although additional research, using a paradigm more suited to infants, will be required to provide more direct evidence for a very early role for the proposed tendencies, evidence for a tendency to attend to intonation in three-year olds, but not adults, does suggest that such a tendency may assist in early language acquisition.

It should be noted that other interpretations of these data are possible. Although the intonational responses of three-year olds, at least, were not simply errors, it is conceivable that children were confused by a test label which incorporated elements of both previously taught labels. Such uncertainty could have resulted in essentially chance responding. An anecdotal description of children's responses in the test situation may partially address that concern. If the test label confused children, then they should have responded more slowly and uncertainly on the test trial. Most children responded as quickly on the test trials as they did on the training trials. A few children did, however, respond by choosing both objects, gave responses such as "it's in the [researcher's] bag," or refused to choose any object. In those cases, the experimenter would say, "if one of these were an 'X,' which one would it be?" If the child did not then choose a single object, the response was recorded as uncodable and that child's score was, as described above, omitted from some analyses (this was true for a total of 5 responses from 4 two-year olds and for 3 responses from 2 three-year olds). Thus, although most children did not seem to be confused by the test label, some children did display evidence of uncertainty. This issue will be more fully addressed in a replication study through a control condition in which children hear an entirely novel label.

General Discussion

The research described here began with the notion that perceptual biases may assist in extracting or identifying words from speech. In the first part of the paper, evidence for the perceptual salience of stressed and final syllables was described: Young children less frequently omitted stressed and final syllables from their productions and, even where unstressed or nonfinal syllables were present, they were less accurate. Tendencies to attend to stressed and final syllables are not, however, sufficient to account for all first productions. A second study explored the possibility that an additional type of attention to stress may also assist children in identifying extactable elements of speech. Children may frequently attend to and extract not only the stressed syllables per se but also the stress *pattern*, or the overall intonation contour, of a sequence of speech. Where subjects were required to identify referents for words on the basis of either stress pattern or phoneme sequence, three-year olds were more likely than adults to make choices based on intonational similarity. Those observations suggest that children are more likely than are

adults to include intonation as an important part of the representation for a word. They are also at least consistent with the view that stress pattern, and possibly other aspects of intonation, may be extracted as part of the representation of a word and may, in some cases, serve to define a word for a child. Thus, it may be that the child is assisted in extracting and identifying first words not only by tendencies to attend to stressed and final syllables, but also by attention to the stress pattern or overall intonation contour of certain sequences.

References

Blasdell, R., & Jensen, P. (1970). Stress and word position as determinants of imitation in first language learners. *Journal of Speech and Hearing Research, 13*, 193-202.

Cole, R. A., & Jakamik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum.

Frumhoff, P., & Newport, E. L. (in preparation). *Perceptual salience and operating principles for language acquisition: The effects of stress and end of word.*

Fueur, H. (1980). Morphological development in Mohawk. *Papers and Reports in Child Language Development, 18*, 25-42.

Gleitman, L. R., & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner, & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3-48). Cambridge, England: Cambridge University Press.

Hayes, J. R., & Clark, H. H. (1970). Experiments on the segmentation of an artificial speech analog. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 221-234). New York: Wiley.

Peters, A. M. (1977). Language learning strategies: Does the whole equal the sum of the parts? *Language, 53*, 560-573.

Peters, A. M. (1983). *The units of language acquisition*. Cambridge, England: Cambridge University Press.

Pye, C. (1983). Mayan telegraphese: Intonational determinants of inflectional development in Quiche Mayan. *Language, 59*, 583-604.

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson, & D. I. Slobin (Eds.), *Studies of child language development* (pp. 175-208). New York: Holt, Rinehart & Winston.

Acknowledgments

I gratefully acknowledge Elissa Newport for ideas and suggestions which contributed substantially to all phases of this research, and for making available recordings analyzed in the first study. I would also like to thank Ann Brown, Dedre Gentner, Michael Kenstowicz, Molly Mack, Ellen Markman and Doug Medin for valuable comments. Ann Garber provided important assistance with transcription. I am grateful to the children, parents, teachers and directors of First United Methodist Daycare and La Petite Academy in Champaign, IL, the Children's Center and Willow Springs Learning Center in Urbana, IL, and Children's Christian Center in Mahomet, IL, as well as the three children who contributed to the corpus for the first study, and their parents. This research was supported by a Cognitive Science/Artificial Intelligence Fellowship from the University of Illinois, by NICHD Training Grant number HD07205, NIH grant number NS16878 to E. Newport and T. Supalla, a dissertation research grant from the Graduate College of the University of Illinois, by Individual National Research Service Award number HD07109 from the National Institute of Child Health and Human Development, and by Sloan Foundation Grant number 83-6-14 to Stanford University. This paper is based, in part, on a dissertation submitted in partial fulfilment of the requirements for the PhD to the University of Illinois, Urbana-Champaign.